

Amendments to the Specification

Please substitute the paragraph beginning on page 2, line 8, with the following paragraph:

Overall, there is increasing evidence that the regulation of normal and aberrant cellular proliferation is not only affected on the transcriptional level, but that also a higher level of regulation is involved, i.e., the organization of chromatin structure through the modification of histone molecules. The determination of the proteins and the molecular mechanisms involved in histone modification will contribute to the understanding of the cellular proliferation program and will thus shed [[led]] light on the mechanisms involved in aberrant proliferation occurring in tumor formation and progression.

Please substitute the paragraphs beginning on page 6, line 17, and ending on page 7, line 2, with the following paragraphs:

In order to achieve this objective, the sequence information of the SET domain was used to obtain the human cDNA homologous to the SET domain genes of *Drosophila* from human CDNA banks. Two cDNAs were obtained which constitute human homologues of *E(z)* and *Su(var)3-9*. The corresponding human genes are referred to as *EZH2* and *SUV39H*. See FIGS. 6A-6C and 7A-7B. In addition, a variant form of *EZH2* was identified which is referred to as *EZH1*. See FIG. 8.

The present invention thus relates to DNA molecules containing a nucleotide sequence coding for a chromatin regulator protein which has a SET-domain, or a partial sequence thereof, characterized in that the nucleotide sequence is that shown in FIGS. 6A-6C

(SEQ ID NO:1), or a partial sequence thereof, or FIGS. 7A-7B (SEQ ID NO:3), or a partial sequence thereof. The DNA molecules, including variants and mutants thereof such as dominant-negative mutants, are also referred to as "genes according to the invention." Two examples of genes according to the invention are designated *EZH2* and *SUV39H*. They were originally referred to as "HEZ-2" and "H3-9," respectively.

Please substitute the paragraph beginning on page 8, line 5, with the following paragraph:

FIGS. 1A-1B ~~[[is]]~~ are an amino acid sequence comparison between *EZH2* (SEQ ID NO:2) and *Drosophila enhancer of zeste (E(z))* (SEQ ID NO:11). The conserved carboxy terminal SET-domain (shaded box) and the Cys-rich region (Cys groups are emphasized) are shown. Percent identity is shown on the right side. The presumed nucleus locating signals are underlined.

Please substitute the paragraph beginning on page 8, line 10, with the following paragraph:

FIGS. 2A-2B ~~[[is]]~~ are an amino acid sequence comparison between the human homologue *SUV39H* (SEQ ID NO:4) and *Drosophila Su(var)3-9* (SEQ ID NO:16). The conserved carboxy terminal SET-domain (shaded box) and the Chromo-domain (darker shaded box) are shown. Percent identity is shown on the right side. The presumed nucleus locating signals are underlined. ~~At the top of the figure~~ FIG. 2A is a diagrammatic summary of the two protein structures which shows that, in the human homologue, 207 amino acids are missing at the N-terminus.

Please substitute the paragraph beginning on page 10, line 3, with the following paragraph:

FIGS. 6A-6C illustrate[[s]] the DNA and amino acid sequences of *EZH2* (SEQ ID NOS:1 and 2, respectively).

Please substitute the paragraph beginning on page 10, line 5, with the following paragraph:

FIGS. 7A-7B illustrate[[s]] the DNA and amino acid sequences of *SUV39H* (SEQ ID NOS:3 and 4, respectively).

Please substitute the paragraph beginning on page 12, line 3, with the following paragraph:

The complete cDNA coding for the human homologue of *E(z)* was designated *EZH2* (SEQ ID NO:1) and the DNA coding for the human homologue of *Su(var)3-9* was designated *SUV39H*(SEQ ID NO:3). All in all, the identity of the amino acids between *Drosophila* and the human proteins amounts to 61% for *EZH2* and 43% for *SUV39H*, whilst the C-terminal SET-domain is very highly conserved (88% for *EZH2* and 53% *SUV39H*). Sequence comparison showed other clear regions of homology, e.g., a cysteine-rich domain in *EZH2* and a Chromo-Box in *SUV39H*. (In *polycomb*, it was shown that the Chromo-Box is the essential domain for the interaction between DNA and chromatin (Messmer, *et al.*, *Genes & Dev.* 6:1241-1254 (1992))[[]]). By contrast, the 207 amino acids which make-up the amino terminal GTP-binding motif of the *Drosophila* protein are absent from the human homologue *SUV39H*. A comparison of the amino acid sequences between *Drosophila* and the human

genes is shown in FIGS. 1A-1B and 2A-2B. Moreover, another cDNA of the SET-domain family known as *MG-44* (see below) also lacks the 5'-end of the *Drosophila* gene.

Please substitute the paragraph beginning on page 18, line 9, with the following paragraph:

Overexpression studies with human SUV39Hmutants indicate a dominant interference with higher-order chromatin organization that, surprisingly, suggests a functional relationship between the SET domain and the distribution of phosphorylated (at serine 10) H3 (Melcher, M., *et al.*, *Mol Cell Biol* 20:3728-41 (2000)). The experiments of the present invention, as shown in the Examples, show that mammalian *SUV39H1*, or other *SUV39H* proteins, are SET domain-dependent, H3-specific histone methyltransferases (HMTases) which selectively methylate lysine 9 of the H3 N-terminus. *See* FIGS. 9A-9B and 10A-10C. Methylation of lysine 9 negatively regulates phosphorylation of serine 10 and reveals a histone code that appears intrinsically linked to the organization of higher-order chromatin.

Please substitute the paragraph beginning on page 24, line 26, with the following paragraph:

Toxicity and therapeutic efficacy of the compounds identified as drug candidates by the methods described above can be determined by standard pharmaceutical procedures, which include conducting cell culture and animal experiments to determine the IC_{50} , $[[LD_{50}]$ LD_{50} and ED_{50} . The data obtained may be used for determining the human dose range, which will also depend on the dosage form (tablets, capsules, aerosol sprays, ampules, etc.) and the administration route (oral, buccal, nasal, parenteral, rectal, etc.). A pharmaceutical composition containing the compound as the active ingredient may be formulated in a

conventional manner using one or more physiologically active carriers and excipients. Methods for making such formulations can be found in manuals, e.g., "Remington Pharmaceutical Sciences."

Please substitute the paragraph beginning on page 28, line 6, with the following paragraph:

The cDNA inserts from recombinant phages were subcloned into the polylinker of pBluescript KS (Stratagene) and sequenced in an automatic sequencer (Applied Biosystems) using the dideoxy method. The complete sequence of at least two independent isolates per gene obtained was determined by primer walking. The sequences were analyzed with the GCG-Software package (University of Wisconsin), and the investigation for homology was carried out using the "Blast and fasta" or "tfasta" network service. The complete sequences of *EZH2* (SEQ ID NO:1) and *SUV39H*(SEQ ID NO:3) are shown in FIGS. 6A-6C and 7A-7B.

Please substitute the originally submitted drawings with the attached 16 sheets of drawings, wherein 15 sheets containing FIGs. 1A, 2A-2B, 3, 4, 5, 6A, 6B, 6C, 7A, 7B, 8, 9A, 9B, 10A, and 10B-10C are labeled Replacement sheets, and one sheet containing FIG. 1B is labeled New sheet.